

Best Available Copy

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-063296

(43)Date of publication of application : 06.03.1998

(51)Int.Cl.

G10L 5/02
G06F 3/16
G06F 17/21
G10L 5/04
H04L 12/54
H04L 12/58
H04M 3/42
H04M 11/00

(21)Application number : 08-235804

(71)Applicant : CANON INC

(22)Date of filing : 20.08.1996

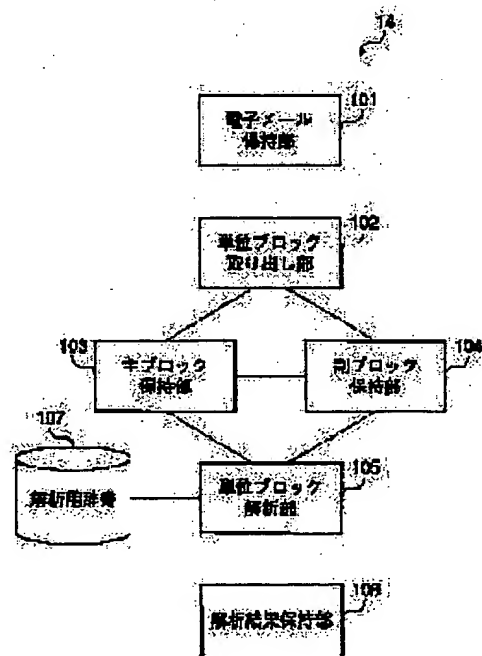
(72)Inventor : AIZAWA MICHIO
KANEKO KAZUE
HIROTA MAKOTO
YAGISAWA TSUYOSHI
FUJITA MINORU

(54) DEVICE AND METHOD FOR NATURAL LANGUAGE PROCESSING

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a device and a method for a natural language processing in which no long time is required for waiting for the processing until a user's command is received in an electronic mail reading device.

SOLUTION: The electronic mail reading device is provided with the natural language processing device 14 and a voice synthesizer. The device 14 takes out one row from the contents of an electronic mail of an electronic mail holding section 101 and copies the contents in an main block holding section 103. Then, analyses are successively made for the characters starting from the leading head of the section 103 till a last character and the result is held in an analysis result holding section 106. If the case exists in which a word spans from the end of the section 103 to a subblock holding section 104, the analysis is made till the word. Based on the analysis result of the section 106, a reading is conducted for every row in a voice synthesizer 15.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-63296

(43)公開日 平成10年(1998)3月6日

(51)Int.Cl. ⁹	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 5/02			G 1 0 L 5/02	G
G 0 6 F 3/16	3 3 0		G 0 6 F 3/16	3 3 0 K
17/21			G 1 0 L 5/04	E
G 1 0 L 5/04			H 0 4 M 3/42	J
H 0 4 L 12/54			11/00	3 0 3

審査請求 未請求 請求項の数 6 F D (全 7 頁) 最終頁に続く

(21)出願番号 特願平8-235804

(22)出願日 平成8年(1996)8月20日

(71)出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72)発明者 相澤 道雄

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72)発明者 金子 和恵

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72)発明者 廣田 誠

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(74)代理人 弁理士 渡部 敏彦

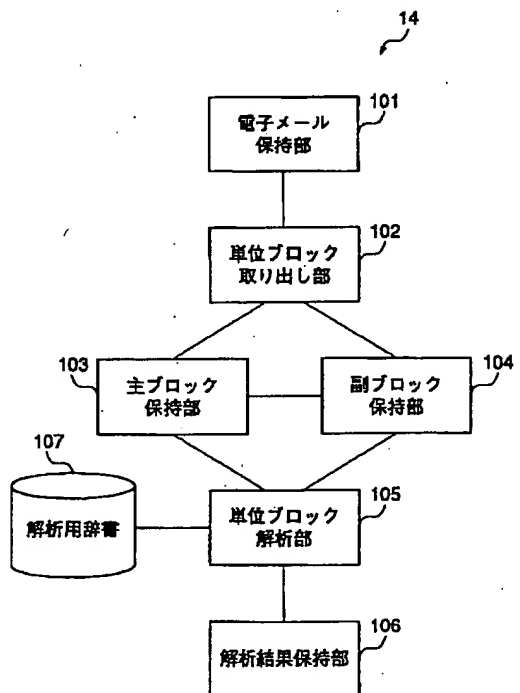
最終頁に続く

(54)【発明の名称】 自然言語処理装置および方法

(57)【要約】

【課題】 電子メール読み上げ装置でユーザのコマンドが受理されるまで長時間待たなくて済む自然言語処理装置および方法を提供する。

【解決手段】 電子メール読み上げ装置は自然言語処理装置14、音声合成装置を備える。自然言語処理装置14では、電子メール保持部101から電子メールの内容を1行取り出し、主ブロック保持部103に複写する。主ブロック保持部103の先頭から順に最後の文字まで解析を行い、その結果を解析結果保持部106に保持する。主ブロック保持部103の末尾から副ブロック保持部104の先頭にまたがった単語がある場合、その単語まで解析する。解析結果保持部106の解析結果を基に、音声合成装置15で1行毎に読み上げを行う。



1

【特許請求の範囲】

【請求項1】 電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理装置において、前記電子メールを保持する電子メール保持手段と、該保持された電子メールから単位ブロックを取り出す単位ブロック取出手段と、該取り出された単位ブロックを保持する単位ブロック保持手段と、該保持された単位ブロックの文字列を読み上げの単位として解析する単位ブロック解析手段とを備えたことを特徴とする自然言語処理装置。

【請求項2】 前記単位ブロック保持手段は、第1の単位ブロックを保持する主単位ブロック保持手段と、前記第1の単位ブロックに後続する第2の単位ブロックを保持する副単位ブロック保持手段とを備えたことを特徴とする請求項1記載の自然言語処理装置。

【請求項3】 前記主単位ブロック保持手段の末尾の文字列と前記副単位ブロック保持手段の先頭の文字列とにまたがって1つの単語が存在している場合、該副単位ブロック保持手段の先頭の文字列を削除すると共に、該先頭の文字列を前記主単位ブロック保持手段の末尾の文字列に連結して前記単位ブロック解析手段により解析することを特徴とする請求項2記載の自然言語処理装置。

【請求項4】 前記単位ブロック解析手段によって使用される解析用辞書を備えたことを特徴とする請求項1記載の自然言語処理装置。

【請求項5】 前記単位ブロックは1行であることを特徴とする請求項1記載の自然言語処理装置。

【請求項6】 電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理方法において、前記電子メールを保持する工程と、該保持された電子メールから単位ブロックを取り出す工程と、該取り出された単位ブロックを保持する工程と、該保持された単位ブロックの文字列を読み上げの単位として解析する工程とを備えたことを特徴とする自然言語処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、電話を使って自分宛の電子メールの内容を確認する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理装置および方法に関する。

【0002】

【従来の技術】近年、ネットワークの発達により電子メールを利用する人が増えている。そして、会議の案内など情報の伝達に電子メールを用いることが多くなった。このため、外出などでコンピュータから長い間離れてい

2

ると、電子メールを読むことができなくなり、大切な情報を知ることができないという問題が生じるようになってきた。

【0003】この問題の対応策として、電話を使って電子メールの内容を確認する電子メール読み上げ装置が提案されている。図6は従来の電子メール読み上げ装置の構成を示すブロック図である。

【0004】ユーザは電子メール読み上げ装置に電話をかけ、プッシュボタンなどを用いて電子メール読み上げ装置にコマンドを送り、自分宛の電子メールの内容を聞く。制御部602は電話603を通じて送られてくるユーザからのコマンドを解釈し、読み上げる電子メールを電子メールデータベース601から取り出す。そして、取り出した電子メールを自然言語処理装置604で解析し、音声合成装置605で音声に変換してユーザの電話606に送信する。

【0005】電子メール読み上げ装置を使うことで、電話があるところならどこでも自分宛の電子メールの内容を確認することができる。

【0006】

【発明が解決しようとする課題】しかしながら、電話を使った電子メール読み上げ装置では、電子メールを読み上げている間、ユーザからのコマンドを受け取りにくい（ほとんど受理できない）ので、ユーザは読み上げと読み上げとの間の休止時間にコマンドを送る必要があった。そのため、電子メール読み上げ装置の読み上げが長く続く場合、ユーザのコマンドが受理されず、長い時間待たされてしまうといったことがあった。

【0007】このように、従来の電子メール読み上げ装置で使われる自然言語処理装置604は一文単位に解析を行い、音声合成装置605はその解析結果を一文単位に読み上げる。例えば、図5のように長い文が現われて読み上げが開始されると、途中でコマンド（例えば、読み上げの中止コマンドなど）を送りたくなくてもコマンドが受理されるまでユーザは長時間待たなければならなかった。

【0008】また、従来の自然言語処理装置では、文の切れめは句点や空行などを利用して判断するが、電子メールには句点や空行のないものが多く、一文が長く判断されることが多い。図7は一文が長く判断されてしまう場合の電子メールの内容を示す図である。この場合、
「中谷さんのレクチャー打ち上げの日程日時：6月13日（木）18：30－場所：西会議室内線：03-5701-4020会費：12日頃に部長以上8000円課長7000円係長6000円その他5000円」といった文字列を一文として判断する。

【0009】この結果、従来の自然言語処理装置を用いた電子メール読み上げ装置では、ユーザのコマンドがなかなか受理されないという問題があった。

【0010】そこで、本発明は、電子メール読み上げ装

置でユーザのコマンドが受理されるまで長時間待たなくて済む自然言語処理装置および方法を提供することを目的とする。

【0011】

【課題を解決するための手段】上記目的を達成するために、本発明の請求項1に係る自然言語処理装置は、電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理装置において、前記電子メールを保持する電子メール保持手段と、該保持された電子メールから単位ブロックを取り出す単位ブロック取出手段と、該取り出された単位ブロックを保持する単位ブロック保持手段と、該保持された単位ブロックの文字列を読み上げの単位として解析する単位ブロック解析手段とを備えたことを特徴とする。

【0012】請求項2に係る自然言語処理装置では、請求項1に係る自然言語処理装置において前記単位ブロック保持手段は、第1の単位ブロックを保持する主単位ブロック保持手段と、前記第1の単位ブロックに後続する第2の単位ブロックを保持する副単位ブロック保持手段とを備えたことを特徴とする。

【0013】請求項3に係る自然言語処理装置では、請求項2に係る自然言語処理装置において前記主単位ブロック保持手段の末尾の文字列と前記副単位ブロック保持手段の先頭の文字列とにまたがって1つの単語が存在している場合、該副単位ブロック保持手段の先頭の文字列を削除すると共に、該先頭の文字列を前記主単位ブロック保持手段の末尾の文字列に連結して前記単位ブロック解析手段により解析することを特徴とする。

【0014】請求項4に係る自然言語処理装置は、請求項1に係る自然言語処理装置において前記単位ブロック解析手段によって使用される解析用辞書を備えたことを特徴とする。

【0015】請求項5に係る自然言語処理装置では、請求項1に係る自然言語処理装置において前記単位ブロックは1行であることを特徴とする。

【0016】請求項6に係る自然言語処理方法は、電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理方法において、前記電子メールを保持する工程と、該保持された電子メールから単位ブロックを取り出す工程と、該取り出された単位ブロックを保持する工程と、該保持された単位ブロックの文字列を読み上げの単位として解析する工程とを備えたことを特徴とする。

【0017】

【発明の実施の形態】本発明の自然言語処理装置および方法の実施の形態について説明する。始めに本実施の形態における自然言語処理装置が適用される電子メール読み上げ装置について説明する。

【0018】図1は実施の形態における電子メール読み上げ装置の構成を示すブロック図である。電子メール読み上げ装置10は、電子メールデータベース11、制御部12、電話13、自然言語処理装置14および音声合成装置15を備える。

【0019】ユーザは電子メール読み上げ装置10に電話をかけてプッシュボタンなどにより電子メール読み上げ装置10にコマンドを送り、自分宛の電子メールの内容を聞く。

【0020】制御部12は、電話16、13を通じて送られてくるユーザからのコマンドを解釈し、読み上げる電子メールを電子メールデータベース11から取り出す。さらに、取り出した電子メールを自然言語処理装置14で解析し、音声合成装置15で音声に変換してユーザの電話16に送信する。

【0021】このように、電子メール読み上げ装置10を使うことで、ユーザは電話があるところならどこでも自分宛の電子メールの内容を確認することができる。

【0022】つぎに、本実施の形態における自然言語処理装置について詳述する。図2は自然言語処理装置の構成を示すブロック図である。

【0023】図において、101は解析される電子メールを保持する電子メール保持部である。102は電子メール保持部101に保持されている電子メールから単位ブロック（例えば、1行など）の文字列を取り出す単位ブロック取り出し部である。103は単位ブロック取り出し部102で取り出された単位ブロックの文字列を保持する主ブロック保持部である。

【0024】104は単位ブロック取り出し部102で取り出された単位ブロックの文字列を保持する副ブロック保持部である。105は単位ブロックの文字列を解析する単位ブロック解析部である。106は単位ブロック解析部105による解析結果を保持する解析結果保持部である。107は単位ブロック解析部105が解析時に使用する解析用辞書である。

【0025】単位ブロック取り出し部102及び単位ブロック解析部105は、後述するフローチャートに対応するプログラムをCPUがメモリから読み出して実行することで実現できる。

【0026】図3は自然言語処理装置の動作処理手順を示すフローチャートである。本実施の形態では、単位ブロックの量が1行である場合を示す。また、図5は自然言語処理装置によって処理される電子メールの内容を示す図である。

【0027】まず、電子メール保持部101から電子メールの内容を1行取り出し、主ブロック保持部103に複写する（ステップS201）。主ブロック保持部103に複写される文字列は、図5の「ユーザ宛に届いた電子メールを格納する電子メール」という1行である。

【0028】主ブロック保持部103の文字列がファイ

ルの終わりを示す記号か否かを判別する(ステップS202)。ファイルの終わりを示す記号である場合、処理を終了する。一方、ファイルの終わりを示す記号でない場合、電子メール保持部101から更に1行取り出し、副ブロック保持部104に複写する(ステップS203)。ここで、副ブロック保持部104に複写される文字列は、「データベースと読み上げる電子メールを選択する電子」という1行である。

【0029】単位ブロック解析部105は、主ブロック保持部103の文字列について自然言語解析処理を行う(ステップS204)。そして、その解析結果を音声合成装置15に送る。さらに、副ブロック保持部104の文字列を主ブロック保持部103に複写し(ステップS205)、ステップS202の処理に戻る。

【0030】図4はステップS204における単位ブロック解析部105の自然言語解析処理手順を示すフローチャートである。ここでは、主ブロック保持部103に保持されている文字列が「ユーザ宛に届いた電子メールを格納する電子メール」、副ブロック保持部104に保持されている文字列が「データベースと読み上げる電子メールを選択する電子」である場合について示す。

【0031】まず、主ブロック保持部103の先頭から順に最後の文字まで解析を行い、その結果を解析結果保持部106に加える(ステップS301)。主ブロック保持部103の末尾から副ブロック保持部104の先頭にまたがった単語がある場合、その単語まで解析する。そして、副ブロック保持部104にまたがった部分を副ブロック保持部104から削除する。

【0032】本実施の形態の場合、解析結果保持部106は、「ユーザ(名詞)宛(接尾辞)に(格助詞)届い(動詞)た(助動詞)電子メール(名詞)を(格助詞)格納する(動詞)電子メールデータベース(名詞)」という解析結果を保持する。ここでは、簡単に単語とその品詞を示したが、音声合成に必要な読みやアクセントなどの情報も解析用辞書107から取り出して解析結果に加えることができる。

【0033】また、単語「電子メールデータベース」のうち、「データベース」が副ブロック保持部104にまたがっているため、その部分を削除し、その結果、副ブロック保持部104は「と読み上げる電子メールを選択する電子」という文字列を保持する。

【0034】そして、副ブロック保持部104の先頭から一単語を解析する(ステップS302)。一単語の解析結果は「と(格助詞)」である。ステップS302で解析した単語が付属語であるか否かを判別する(ステップS303)。付属語でない場合、処理を終了し、付属語である場合、解析した単語(付属語)を解析結果保持部106に加え、(ステップS304)、副ブロック保持部104からその一単語を削除する。この後、ステップS302の処理に戻る。

【0035】ここまでの結果、解析結果保持部106は「ユーザ(名詞)宛(接尾辞)に(格助詞)届い(動詞)た(助動詞)電子メール(名詞)を(格助詞)格納する(動詞)電子メールデータベース(名詞)と(格助詞)」を保持し、副ブロック保持部104は「読み上げる電子メールを選択する電子」を保持する。

【0036】本実施の形態の場合、ステップS302で次に解析される単語が「読み(動詞)」であり付属語でないため、ステップS303で処理を終了する。

【0037】最終的に、解析結果保持部106は「ユーザ(名詞)宛(接尾辞)に(格助詞)届い(動詞)た(助動詞)電子メール(名詞)を(格助詞)格納する(動詞)電子メールデータベース(名詞)と(格助詞)」を保持することになり、副ブロック保持部104は「読み上げる電子メールを選択する電子」を保持することとなる。

【0038】そして、電子メール読み上げ装置10の音声合成装置15は、自然言語処理装置14の解析結果である単位ブロック(1行)毎に読み上げを行う。したがって、読み上げが開始された後、ユーザは長時間待つことなく単位ブロック(1行)毎の読み上げと読み上げとの間にコマンド(例えば、読み上げの中止コマンドなど)を送信してコマンドを受理して貰うことができる。

【0039】尚、本実施の形態では、単位ブロックの量を1行としたが、例えば76バイトのように文字の量で決めてもよい。

【0040】また、ステップS204で主ブロック保持部103の文字列を解析し終わっても、付属語が続く場合、続けて副ブロック保持部104の文字列を解析するが、主ブロック保持部103の文字列を解析し終わったら、次に付属語が続く場合でもそこで解析を終了するようにしてもよい。

【0041】さらに、本実施の形態では、単位ブロックの量を固定としたが、電子メール読み上げ装置内の制御部などが動的に単位ブロックの量を変更するようにしてもよい。例えば、電子メールの先頭部分では単位ブロックの量を少なく設定し、後半部分では単位ブロックの量を多く設定するようにしてもよい。

【0042】また、本実施の形態では単位ブロックの量を自然言語処理装置が決定していたが、ユーザがプッシュボタンなどを用いて指定するようにしてもよい。

【0043】さらに、本発明は複数の機器から構成されるシステムに適用してもよいし、1つの機器からなる装置に適用してもよい。また、本発明はシステムあるいは装置にプログラムを供給することによって達成される場合にも適用できることはいうまでもない。この場合、本発明を達成するためのソフトウェアによって表されるプログラムを格納した記憶媒体をシステムあるいは装置に読み出すことによってそのシステムあるいは装置が本発明の効果を享受することが可能となる。

【 0044 】

【 発明の効果 】 本発明の請求項1に係る自然言語処理装置によれば、電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する際、電子メール保持手段により前記電子メールを保持し、単位ブロック取出手段により該保持された電子メールから単位ブロックを取り出し、単位ブロック保持手段により該取り出された単位ブロックを保持し、単位ブロック解析手段により該保持された単位ブロックの文字列を読み上げの単位として解析するので、単位ブロック毎に読み上げを行い、その読み上げと読み上げとの間にコマンドを送ることができ、ユーザのコマンドが受理されるまで長時間待たなくて済む。

【 0045 】 請求項2に係る自然言語処理装置によれば、前記単位ブロック保持手段は、第1の単位ブロックを保持する主単位ブロック保持手段と、前記第1の単位ブロックに後続する第2の単位ブロックを保持する副単位ブロック保持手段とを備えたので、自然言語処理装置によって解析されて読み上げられる単位ブロックの区切り部分を、後続部分を参照して処理を行うことができ

る。

【 0046 】 請求項3に係る自然言語処理装置によれば、前記主単位ブロック保持手段の末尾の文字列と前記副単位ブロック保持手段の先頭の文字列とにまたがって1つの単語が存在している場合、該副単位ブロック保持手段の先頭の文字列を削除すると共に、該先頭の文字列を前記主単位ブロック保持手段の末尾の文字列に連結して前記単位ブロック解析手段により解析するので、自然言語処理装置によって解析されて読み上げられる単位ブロックを単語の境界で区切ることができ、ユーザにとって聞き易くなる。

【 0047 】 請求項4に係る自然言語処理装置によれば、前記単位ブロック解析手段によって使用される解析用辞書を備えたので、多種多様な解析を行うことができる。

【 0048 】 請求項5に係る自然言語処理装置によれば、前記単位ブロックは1行であるので、1行単位に読み上げを行うことができる。

【 0049 】 請求項6に係る自然言語処理方法によれば、電子メールの内容を音声出力する電子メール読み上げ装置に使用され、前記読み上げられる電子メールの内容を解析する自然言語処理方法において、前記電子メールを保持する工程と、該保持された電子メールから単位ブロックを取り出す工程と、該取り出された単位ブロックを保持する工程と、該保持された単位ブロックの文字列を読み上げの単位として解析する工程とを備えたので、単位ブロック毎に読み上げを行い、その読み上げと読み上げとの間にコマンドを送ることができ、ユーザのコマンドが受理されるまで長時間待たなくて済む。

【 図面の簡単な説明 】

【 図1 】 実施の形態における電子メール読み上げ装置の構成を示すブロック図である。

【 図2 】 自然言語処理装置の構成を示すブロック図である。

【 図3 】 自然言語処理装置の動作処理手順を示すフローチャートである。

【 図4 】 ステップS204における単位ブロック解析部105の自然言語解析処理手順を示すフローチャートである。

【 図5 】 自然言語処理装置によって処理される電子メールの内容を示す図である。

【 図6 】 従来の電子メール読み上げ装置の構成を示すブロック図である。

【 図7 】 一文が長く判断されてしまう場合の電子メールの内容を示す図である。

【 符号の説明 】

10 電子メール読み上げ装置

14 自然言語処理装置

15 音声合成装置

101 電子メール保持部

102 単位ブロック取り出し部

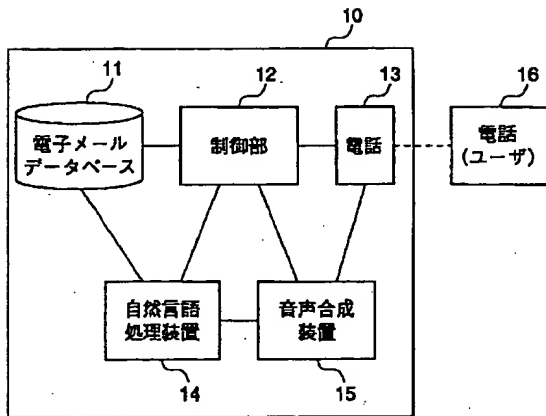
103 主ブロック保持部

104 副ブロック保持部

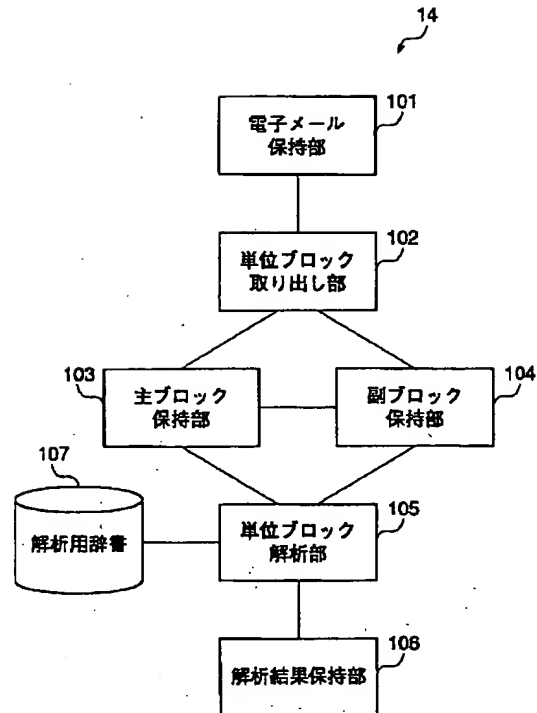
105 単位ブロック解析部

107 解析用辞書

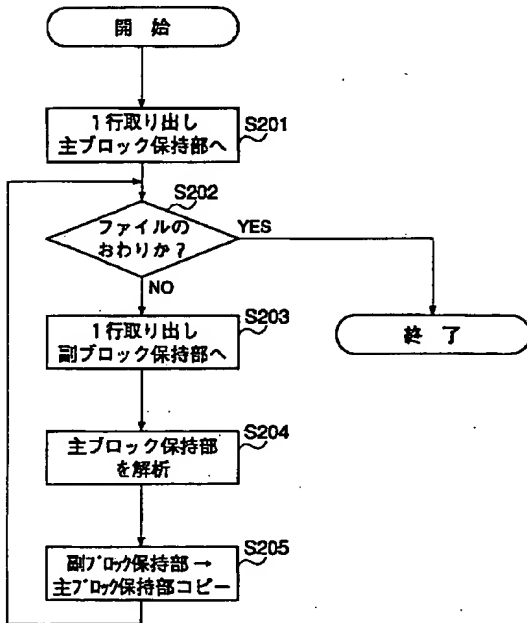
【 図1 】



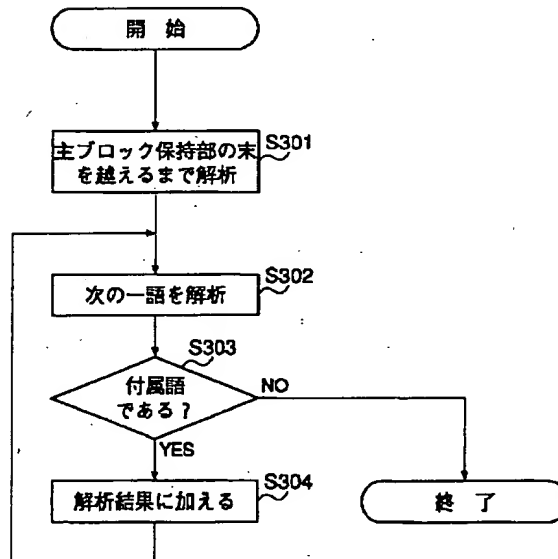
【 図2 】



【 図3 】



【 図4 】



【 図5 】

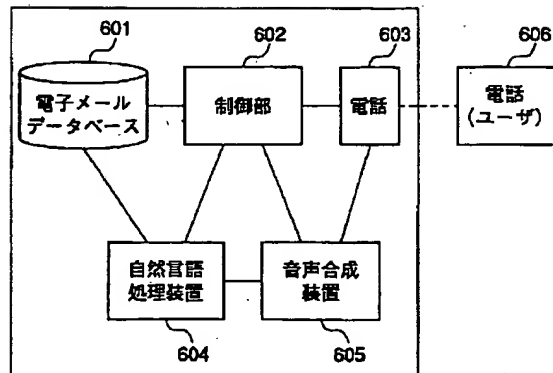
TO: minoro
Cc: yag611, kaneki, keicho, hiru
From: aisawa

ユーザ宛に届いた電子メールを格納する電子メールデータベースと読み上げる電子メールを選択する電子メール選択部と電子メールの選択に利用する条件を保持する選択条件保持部と電子メールを音声に変換する音声合成部とシステムを利用しているユーザがだれであるかを確認するユーザ確認部とを備えることを特徴とする電子メール読み上げ装置。

請求項です。

-- aisawa

【 図6 】



【 図7 】

TO: minoro, yag611, kaneki, keicho, hiru
Cc: ika
From: ika

中谷さんのレクチャー打ち上げの日程

日時: 6月13日 (木) 18:30-

場所: 西会議室

内線: 03-5701-4020

会費: 12日頃に

部長以上 8000円

課長 7000円

係長 6000円

その他 5000円

当日は17時50分頃出ます。

直接行かれる方は地図があるので、お知らせ下さい。

ika

フロント ページの続き

(51) Int. Cl. 6	識別記号	庁内整理番号	FI	技術表示箇所
H04L 12/58			G06F 15/20	568A
H04M 3/42		9744-5K	H04L 11/20	101B
11/00	303			

(72)発明者 八木沢 津義
東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72)発明者 藤田 稔
東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内